Case Study for Breakout Session #1

A sponsor is developing a new drug (Drug X) and has found that some individuals experience hepatotoxicity. They want to identify biomarkers in peripheral blood white cells of those patients who are sensitive to Drug X. They undertake an investigative study utilizing Affymetrix GeneChip® arrays that monitor thousands of human genes and test several hundred human blood samples, including some from patients that are known to be sensitive to Drug X.

Utilizing a platform that monitors the expression level of thousands of genes at one time allows normalization of expression ratios (i.e., fluorescent signals) based on multiple genes. Those genes can be selected to represent different expression levels lending themselves to the generation of standard curves. In this instance, data analysis identifies 10 dysregulated genes in this subpopulation of patients that statistical analysis suggests might be good biomarkers of Drug X sensitivity.

Questions:

1. How many biological samples from sensitive and non-sensitive patients are needed to derive a gene list? How could one calculate the potential number needed?
2. What analysis approaches might one utilize? How can one determine if the appropriate approach was used?

To test these genes in clinical trials, the sponsor will move from an array approach to a testing system that is more adaptable to many clinical laboratories and that can more efficiently monitor this small gene set. The sponsor selects a TaqMan® RT-PCR platform. However, it now becomes critical that an appropriate gene is selected as a comparator (i.e., housekeeping gene) for use in standardizing all samples. Using the Affymetrix GeneChip® data generated from the several hundred human blood samples as described above, their analysis leads them to select "Gene A" whose expression level shows no statistically significant difference between any populations within that dataset.

Questions:

1. How many biological samples would be needed to approach this analysis given the inter-individual variability, differences in metabolic states, etc.?
2. How critical would it be to have data from patients with underlying disease vs. normal individuals?
3. What data analysis approaches might one utilize?

The sponsor selects regions for amplification of the 10 test genes and Gene A thought to be relatively resistant to RNA degradation to minimize, as much as possible, the impact of RNA quality on the assay. They distribute these probes to laboratories proficient in running TaqMan-based assays. After 1000 samples are analyzed, they discover that Gene A is not a good comparator as it expression level appears to be modified in patients with active arthritis, a target population for Drug X treatment.

Questions:

1. How might the sponsor find a better comparator gene?
2. Assuming a better gene is selected, would they be able to reanalyze the prospective samples or would they be required to re-start the clinical trial to process only prospective samples?